#### Raw Sequence Listing

An 180#//

03/29/93 45:40:13

S543.raw

Patent Application US/07/923,692

1 SEQUENCE LISTING ( 2 3 (1) GENERAL INFORMATION: 4 5 (i) APPLICANT: Donson, Jon 6 Dawson, William 0. 7 Grantham, George L. \* In addition, this 8 Turpen, Thomas H. 9 Turpen, Ann Myers 10 Garger, Stephen J. 11 Grill, Laurence K. 12 TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS SUbmiles 13 (ii) 14 (iii) 15 NUMBER OF SEQUENCES: 11 16 text. Prease 17 (iv) CORRESPONDENCE ADDRESS: 18 ADDRESSEE: Limbach & Limbach ensure that 19 STREET: 2001 Ferry Building 20 (C) CITY: San Francisco 21 (D) STATE: CAL ZIP: 94111 22 (F) 23 (v) 24 COMPUTER READABLE FORM: 25 (A) MEDIUM TYPE: Floppy disk 26 COMPUTER: IBM PC compatible (B) OPERATING SYSTEM: PC-DOS/MS-DOS 27 (C) 28 (D) SOFTWARE: Patent in Release #1.0, Version #1.25 29 30 (vi) CURRENT APPLICATION DATA: 31 APPLICATION NUMBER: (A) (B) FILING DATE: 32 33 (C) CLASSIFICATION: 34 (vii) PRIOR APPLICATION DATA: (A) APPLICATION NUMBER: US 600,244 35 36 (B) FILING DATE: 22-OCT-1990 37 PRIOR APPLICATION DATA: 38 (vii) 39 APPLICATION NUMBER: US 641,617 40 (B) FILING DATE: 16-JAN-1991 41 PRIOR APPLICATION DATA: 42 (vii) 43 APPLICATION NUMBER: US 310,881 44 FILING DATE: 17-FEB-1989 45 PRIOR APPLICATION DATA: 46 (vii) 47 (A) APPLICATION NUMBER: US 160,766 48 (B) FILING DATE: 26-FEB-1988 49 50 (vii) PRIOR APPLICATION DATA: 51 (A) APPLICATION NUMBER: US 160,771 52 (B) FILING DATE: 26-FEB-1988

```
53
 54
     (vii)
               PRIOR APPLICATION DATA:
 55
                (A) APPLICATION NUMBER: US 347,637
 56
               (B) FILING DATE: 05-MAY-1989
 57
 58
     (vii)
               PRIOR APPLICATION DATA:
 59
                (A) APPLICATION NUMBER: US 363,138
 60
                   FILING DATE: 08-JUN-1989
 61
 62
     (vii)
               PRIOR APPLICATION DATA:
 63
                (A) APPLICATION NUMBER: US 219,279
 64
                (B) FILING DATE: 15-JUL-1988
 65
 66
               ATTORNEY/AGENT INFORMATION:
     (viii)
 67
                (A) NAME: Halluin, Albert P.
 68
                    REGISTRATION NUMBER: 28,957
 69
                (C) REFERENCE/DOCKET NUMBER: BIOG-20121 USA
 70
 71
               TELECOMMUNICATION INFORMATION:
      (ix)
 72
                    TELEPHONE: 415-433-4150
 73
                (B)
                    TELEFAX: 415-433-8716
 74
 75
 76
     (2) INFORMATION FOR SEQ ID NO: 1:
 77
 78
       (i)
               SEQUENCE CHARACTERISTICS:
 79
               (A)
                   LENGTH: 4 amino acids
 80
                (B)
                    TYPE: amino acid
 81
               (D)
                   TOPOLOGY: linear
 82
 83
      (ii)
               MOLECULE TYPE: peptide
 84
 85
     (iii)
               HYPOTHETICAL: NO
 86
 87
      (iv)
               ANTI-SENSE: NO
 88
 89
               SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 90
 91
          Pro Xaa Gly Pro
 92
          1
 93
 94
     (2) INFORMATION FOR SEQ ID NO: 2:
 95
       (i)
 96
               SEQUENCE CHARACTERISTICS:
 97
                (A) LENGTH: 13 base pairs
 98
                (B) TYPE: nucleic acid
                    STRANDEDNESS: single
 99
                (C)
100
                (D) TOPOLOGY: linear
101
102
      (ii)
               MOLECULE TYPE: DNA (genomic)
103
104
    (iii)
               HYPOTHETICAL: NO
```

### Raw Sequence Listing

03/29/93 15:40:16 S543.raw

105															
105 106	(iv)	NAME CENTER AND													
107	(10)	ANTI-SENSE: NO													
108	(xi)	CENTENCE DECEDIDATON, GEO ID NO. 1.													
109	(XI)	SEQUENCE DESCRIPTION: SEQ ID NO: 2:													
110	GGGTACC														
111	GGGIACC	166 600	13												
112															
113															
114	(2) IN	FORMATION FOR SEQ ID NO: 3:													
115	(2) 111.	FORMATION FOR BEQ ID NO: 3:													
116	(i)	SEQUENCE CHARACTERISTICS:													
117	(1)	(A) LENGTH: 886 base pairs													
118		<u>-</u>													
119	(B) TYPE: nucleic acid (C) STRANDEDNESS: single														
120															
121		(b) TOPOLOGI: Timear													
122	(ii)	MOLECULE TYPE: DNA (genomic)													
123	(11)	MODECODE TIPE: DAR (GeNOMIC)													
124	(iii)	HYPOTHETICAL: NO													
125	(111)	HIFOIREITCAL: NO													
126	(iv)	ANTI-SENSE: NO													
127	(10)	MII-DENDE. NO													
128	(vi)	ORIGINAL SOURCE:													
129	( • = /	(A) ORGANISM: Chinese cucumber													
130		(A) CROMIDM. CHINESE CUCMIDEL													
131	(vii)	IMMEDIATE SOURCE:													
132	( /	(B) CLONE: alpha-trichosanthin													
133		(b) Choke. alpha-ciichobanchin													
134	(ix)	FEATURE:													
135	(11)	(A) NAME/KEY: CDS (B) LOCATION: 8877													
136		(B) LOCATION: 8877													
137		(D) LOCATION. UU//													
138	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 3:													
139	(,	begonice sescritizati seg is no. J.													
140	CTCGAGG	ATG ATC AGA TTC TTA GTC CTC TCT TTG CTA ATT CTC ACC CTC	49												
141	0.0000		4,5												
142		Met Ile Arg Phe Leu Val Leu Ser Leu Leu Ile Leu Thr Leu													
143		1 5 10													
144															
145	TTC CTA	ACA ACT CCT GCT GTG GAG GGC GAT GTT AGC TTC CGT TTA TCA	97												
146			,												
147	Phe Leu	Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe Arg Leu Ser													
148	15	20 25 30													
149															
150	GGT GCA	ACA AGC AGT TCC TAT GGA GTT TTC ATT TCA AAT CTG AGA AAA	145												
151		The second secon													
152	Gly Ala	Thr Ser Ser Ser Tyr Gly Val Phe Ile Ser Asn Leu Arg Lys													
153	-	35 40 45													
154															
155	GCT CTT	CCA AAT GAA AGG AAA CTG TAC GAT ATC CCT CTG TTA CGT TCC	193												
156															

### Raw Sequence Listing

03/29/93 15:40:17 S543.raw

157 158 159	Ala	Leu	Pro	Asn 50	Glu	Arg	Lys	Leu	Tyr 55	Asp	Ile	Pro	Leu	Leu 60	Arg	Ser	
160 161	TCT	CTT	CCA	GGT	TCT	CAA	CGC	TAC	GCA	TTG	ATC	CAT	CTC	ACA	AAT	TAC	241
162 163 164	Ser	Leu	Pro 65	Gly	Ser	Gln	Arg	<b>Tyr</b> 70	Ala	Leu	Ile	His	Leu 75	Thr	Asn	Tyr	
165 166	GCC	GAT	GAA	ACC	ATT	TCA	GTG	GCC	ATA	GAC	GTA	ACG	AAC	GTC	TAT	ATT	289
167 168 169	Ala	qaA 08	Glu	Thr	Ile	Ser	Val 85	Ala	Ile	Asp	Val	Thr 90	Asn	Val	Tyr	Ile	
170 171	ATG	GGA	TAT	CGC	GCT	GGC	GAT	ACA	TCC	TAT	TTT	TTC	AAC	GAG	GCT	TCT	337
172 173 174	Met 95	Gly	Tyr	Arg	Ala	Gly 100	Asp	Thr	Ser	Tyr	Phe 105	Phe	Asn	Glu	Ala	Ser 110	
175 176	GCA	ACA	GAA	GCT	GCA	AAA	TAT	GTA	TTC	AAA	GAC	GCT	ATG	CGA	AAA	GTT	385
177 178 179	Ala	Thr	Glu	Ala	Ala 115	Lys	Tyr	Val	Phe	Lys 120	Asp	Ala	Met	Arg	Lys 125	Val	
180 181																GGC	433
182 183 184	Thr	Leu	Pro	Tyr 130	Ser	Gly	Asn	Tyr	Glu 135	Arg	Leu	Gln	Thr	Ala 140	Ala	Gly	
185 186	AAA	ATA	AGG	GAA	AAT	ATT	CCG	CTT	GGA	CTC	CCA	GCT	TTG	GAC	AGT	GCC	481
187 188 189	Lys	Ile	Arg 145	Glu	Asn	Ile	Pro	Leu 150	Gly	Leu	Pro	Ala	Leu 155	Asp	Ser	Ala	
190 191	ATT	ACC	ACT	TTG	TTT	TAC	TAC	AAC	GCC	AAT	TCT	GCT	GCG	TCG	GCA	CTT	529
192 193 194	Ile	Thr 160	Thr	Leu	Phe	Tyr	Tyr 165	Asn	Ala	Asn	Ser	Ala 170	Ala	Ser	Ala	Leu	
195 196	ATG	GTA	CTC	ATT	CAG	TCG	ACG	TCT	GAG	GCT	GCG	AGG	TAT	AAA	TTT	ATT	577
197 198 199	Met 175	Val	Leu	Ile	Gln	Ser 180	Thr	Ser	Glu	Ala 185	Ala	Arg	Tyr	Lys	Phe	Ile 190	
200 201	GAG	CAA	CAA	ATT	GGG	AAG	CGC	GTT	GAC	AAA	ACC	TTC	CTA	CCA	AGT	TTA	625
202 203 204	Glu	Gln	Gln	Ile	Gly 195	Ľув	Arg	Val	Asp	Lys 200	Thr	Phe	Leu	Pro	Ser 205	Leu	
205 206						GAA											673
207 208	Ala	Ile	Ile	Ser 210	Leu	Glu	Asn	Ser	Trp 215	Ser	Ala	Leu	Ser	Lys 220	Gln	Ile	

# Raw Sequence Listing

03/29/93 15:40:19 S543.raw

209																	
210	CAG	ATA	GCG	AGT	ACT	AAT	AAT	GGA	CAG	TTT	GAA	ACT	CCT	GTT	GTG	CTT	721
211	<b>63</b>	<b>-</b> 1 -		<b>~</b>	<b>~</b> 1	•				_,			_	•		_	
212 213	GIN	TTE	Ala	ser	Inr	ABN	Asn	_	GIN	Pne	GIU	Thr		vaı	Val	Leu	
213			225					230					235				
215	2012	220	aam	<b>733</b>	330	<b>a.</b>	<b>a</b> as	ama	3 m.c	3.003			-	~~		~~-	
215	AIA	WWI	GCT	CAA	AAC	CAA	CGA	GTC	ATG	ATA	ACC	AAT	GTT	GAT	GCT	GGA	769
217	Tla	A an	77.	~1 <b>~</b>	2	<b>71</b> m	7	171	W-4	<b>T1</b> -	mb	3	77_7	<b>3</b>		<b>a</b> 1	
217	110	240	Ala	GIII	WRII	GIII	245	Val	Mec	TIE	Thr		var	Авр	АТА	GIA	
219		240					243					250					
220	CTT	GT A	» CC	TCC	220	አጥሮ	aca	ጥጥረ	CITIC!	CITIC	2 2 77	CCA	220	3 3 M	3 m/a	GCA	017
221	GII	GIA	ACC	100	MAC	AIC	GCG	116	CIG	CIG	AAI	CGA	AAC	AAT	ATG	GCA	817
222	17 n 1	17=1	Thr	G0~	yan	Tla	. הוג	T 011	T 011	T 011	2 an	7 ~~	<b>3</b> a m	3 am	W-+	77-	
223	255	Val	1111	Per	VDII	260	ATG	пеп	пеп	Ten	265	Arg	ASII	ASII	met	270	
224	233					200					200					270	
225	GCC	AΤC	GAT	GAC	CAT	CTT	ССТ	እጥር	a Ca	CNG	» CC	mmm.	CCN	ጥረጥ	CCA	N CITT	865
226	900	AIG	GAI	GAC	GAI	GII	CCI	AIG	ACA	CAG	AGC	111	GGA	161	GGA	AGI	865
227	Δla	Met	Asp	Agn	λan	Val	Pro	Mot	Thr	Gl n	Sar	Dho	Gl <sub>32</sub>	Caro	Gl <sub>W</sub>	Cor	
228		1100	***DP	nop	275		110	Mec	1111	280	Der	F 116	GIY	Cyb	285	per	
229						•				200					200		
230	TAT	GCT	ATT	TAG'	raac:	rcg <i>i</i>	AG										886
231							_										
232	Tyr	Ala	Ile														
233	•			29	)												
234																	
235																	
235 236	(2)	IN	FORM	ATIO	N FOI	R SE(	) ID	NO:	<b>1</b> :								
	(2)	IN	FORM	ATIO	n Foi	R SE(	) ID	NO:4	1:								
236	•	INI L)			N FOI	•	_										
236 237	•		SI	EQUE		CHARI	- ACTEI	RIST	cs:	ids							
236 237 238	•		S1	EQUEI	NCE (	CHARI	ACTEI 289 a	RIST:	cs:	ids							
236 237 238 239	•		SI (1	EQUEI A) I	NCE (	CHARI CH: 2	ACTEI 289 a ino a	RIST: amino acid	cs:	Lds							
236 237 238 239 240 241 242	(1	L)	SI (1	EQUEI A) I	NCE ( LENG! LYPE:	CHARI CH: 2	ACTEI 289 a ino a	RIST: amino acid	cs:	ids							
236 237 238 239 240 241 242 243	•	L)	S1 (1 (1	EQUEI A) I B) I	NCE ( LENG! LYPE:	CHARI TH: 2 : am: LOGY:	ACTEI 289 a ino a : lir	RIST: amino acid near	ICS:	ids							
236 237 238 239 240 241 242 243 244	(i	L) L)	S1 (1 (1 (1	EQUEI A) I B) I D) I	NCE ( LENG! TYPE: TOPO!	CHARI TH: 2 : am: LOGY:	ACTEI 289 a ino a : lir	RIST: amino acid near	ICS: o ac:								
236 237 238 239 240 241 242 243 244 245	(i:	L) L)	SI (1 (1 (1 M(	EQUEI	NCE ( LENGI TYPE: TOPOI ULE I	CHARI TH: 2 : am: LOGY: TYPE:	ACTEI 289 a ino a : li: : pro	RIST:	ICS: o ac: n	ID 1							
236 237 238 239 240 241 242 243 244 245 246	(i:	L) L)	SI (1 (1 (1 M(	EQUEI	NCE ( LENGI TYPE: TOPOI ULE I	CHARI TH: 2 : am: LOGY: TYPE:	ACTEI 289 a ino a : li: : pro	RIST:	ICS: o ac: n	ID 1							
236 237 238 239 240 241 242 243 244 245 246 247	(i.i.) (x:i	L) L)	SI (1 (1 (1 M(	EQUEI	NCE ( LENGI TYPE: TOPOI ULE I	CHARI TH: 2 : am: LOGY: TYPE:	ACTEI 289 a ino a : li: : pro	RIST:	ICS: o ac: n	ID 1							
236 237 238 239 240 241 242 243 244 245 246 247 248	(ii (xi	L) L)	S1 (1 (1 (1	EQUEI	NCE ( LENGT TYPE TOPOI  NCE I	CHARI TH: 2 : am: LOGY: TYPE:	ACTEI 289 a ino a : li: : pro	RIST:	ICS: o ac: n	ID 1			Thr	Leu		Leu	
236 237 238 239 240 241 242 243 244 245 246 247 248 249	(i.i.) (x:i	L) L)	SI (1 (1 (1 M(	EQUEI	NCE ( LENGI TYPE: TOPOI ULE I	CHARI TH: 2 : am: LOGY: TYPE:	ACTEI 289 a ino a : li: : pro	RIST:	ICS: o ac: n	ID 1			Thr	Leu	Phe 15	Leu	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250	(i: (x: Met	i) i) [][]=	SI (1) (1) (1) M(1) SI	EQUEI DLECT EQUEI Phe	NCE ( LENGT TYPE: TOPOI  NCE I  Leu 5	CHARI TH: Z : am: LOGY: TYPE: DESCI	ACTER 289 a ino a : lin : pro RIPT:	RIST: amino acid hear otein ION:	SEQ	ID 1	Ile	Leu			15		
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251	(i: (x: Met	i) i) [][]=	SI (1 (1 (1 M(	EQUEI A) 1 B) 5 DLECT EQUEI Phe	NCE ( LENGT TYPE: TOPOI  NCE I  Leu 5	CHARI TH: Z : am: LOGY: TYPE: DESCI	ACTER 289 a ino a : lin : pro RIPT:	RIST: amino acid hear otein ION:	SEQ Leu	ID 1	Ile	Leu		Ser	15		
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252	(i: (x: Met	i) i) [][]=	SI (1) (1) (1) M(1) SI	EQUEI DLECT EQUEI Phe	NCE ( LENGT TYPE: TOPOI  NCE I  Leu 5	CHARI TH: Z : am: LOGY: TYPE: DESCI	ACTER 289 a ino a : lin : pro RIPT:	RIST: amino acid hear otein ION:	SEQ	ID 1	Ile	Leu			15		
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253	(iii (xi	i) i) (Thr	SI (1) (1) (1) M(1) SI Arg	EQUED  COLECT  EQUED  Phe  Ala  20	NCE (LENGTOPO)  LEU  Val	CHARI TH: Z CHARI	ACTER 289 s ino s : lin : pro RIPT: Leu Gly	RIST: amino acid near oteir ION: Ser	SEQ Leu	ID 1 Leu 10 Ser	Ile Phe	Leu Arg	Leu	Ser 30	15 Gly	Ala	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254	(iii (xi	i) i) (Thr	SI (1) (1) (1) M(1) SI Arg	EQUED  COLECT  EQUED  Phe  Ala  20	NCE (LENGTOPO)  LEU  Val	CHARI TH: Z CHARI	ACTER 289 s ino s : lin : pro RIPT: Leu Gly	RIST: amino acid near oteir ION: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser	Ile Phe	Leu Arg	Leu	Ser 30	15 Gly	Ala	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 253 254 255	(iii (xi	i) i) (Thr	SI (1) (1) (1) M(1) SI Arg	EQUED  COLECT  EQUED  Phe  Ala  20	NCE (LENGTOPO)  LEU  Val	CHARI TH: Z CHARI	ACTER 289 s ino s : lin : pro RIPT: Leu Gly	RIST: amino acid near oteir ION: Ser Asp	SEQ Leu	ID 1 Leu 10 Ser	Ile Phe	Leu Arg	Leu	Ser 30	15 Gly	Ala	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 253 254 255 256	(iii (xi  Met  1 Thr	i) i) Thr	SI (1 (1 M) SI Arg Pro	EQUENTAL STATE OF THE PROPERTY	Leu  Leu  Tyr  Tyr	CHARI TH: Z : am : COGY: CYPE: DESCI Val Val	ACTER 289 a ino a : lin : pro RIPT: Leu Gly Val	RIST: amino acid hear otein ION: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 253 254 255 257	(iii (xi  Met  1 Thr	i) i) Thr Ser	SI (1) (1) (1) M(1) SI Arg	EQUENTAL STATE OF THE PROPERTY	Leu  Leu  Tyr  Tyr	CHARI TH: Z : am : COGY: CYPE: DESCI Val Val	ACTER 289 8 ino 8 : lin : pro RIPT: Leu Gly Val	RIST: amino acid hear otein ION: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
236 237 238 239 240 241 242 243 244 245 246 247 248 250 251 253 254 256 257 258	(iii (xi  Met  1 Thr	i) i) Thr	SI (1 (1 M) SI Arg Pro	EQUENTAL STATE OF THE PROPERTY	Leu  Leu  Tyr  Tyr	CHARI TH: Z : am : COGY: CYPE: DESCI Val Val	ACTER 289 a ino a : lin : pro RIPT: Leu Gly Val	RIST: amino acid hear otein ION: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 253 254 255 257	(iii (xi  Met  1  Thr	Thr Ser Asn 50	SI (1 (1 M) SI Arg Pro	EQUED SOLECTION OF THE PROPERTY OF THE PROPERT	Leu Type: Topoi ULE 1 NCE I Leu 5 Val	CHARI TH: 2 : am: LOGY: TYPE: DESCI Val Glu Glu	ACTER 289 8 ino 8 : lin : pro RIPT: Leu Gly Val Tyr 55	RIST: amino acid near otein ION: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser Ser	Ile Phe Asn Leu	Leu Arg Leu Leu 60	Leu Arg 45 Arg	Ser 30 Lys Ser	15 Gly Ala Ser	Ala Leu Leu	

### Raw Sequence Listing

03/29/93 15:40:20 S543.raw

261	65					70					75					80
262	<b>a</b> 1	m1	<b>~1</b> -		**- 7			_			_		_			
263 264	GIU	Thr	TTG	ser		Ala	тте	Asp	val		Asn	Val	Tyr	Ile		GIÀ
265					85					90					95	
266	Тчг	7~~	710	<b>a</b> 1	N an	Th-	Com	The sea	Dh.	Dh.	3	<b>a</b> 1		<b>a</b>		ml <sub>a as</sub>
267	TYT	ALG	AIG	100	waħ	Thr	Ser	ıyr		Pne	ABN	GIU	АТЯ		ATA	Thr
268				100					105					110		
269	G1.,	λΙο	77-	Tara	П	Val	Dho	Tara	3	21.	<b>W</b> - <b>b</b>	3	<b>T</b>	**- 7	m1	<b>*</b>
270	GIU	ATQ	115	пув	TÄT	Val	FIIG	120	мвр	ATA	met	Arg		vaı	Thr	Leu
271			113					120					125			
272	Pro	Tvr	Ser	Glv	λan	Tyr	Glu	Ara	T.011	Gln	Thr	<b>Δ</b> Ι =	λla	Glar	Tara	TIO
273		130		<b></b> 1		-3-	135	9		<b>U</b>	1111	140	AIG	GLY	шyв	116
274												110				
275	Ara	Glu	Asn	Ile	Pro	Leu	Glv	Len	Pro	Ala	T.em	Δan	Ser	Δla	τlα	Thr
276	145					150	<b>-</b> _1				155	p	DCL	niu	110	160
277																100
278	Thr	Leu	Phe	Tvr	Tvr	Asn	Ala	Asn	Ser	Ala	Ala	Ser	Ala	Leu	Met	Val
279				•	165					170					175	
280																
281	Leu	Ile	Gln	Ser	Thr	Ser	Glu	Ala	Ala	Arq	Tyr	Lvs	Phe	Ile	Glu	Gln
282				180					185	_	•	•		190		
283																
284	Gln	Ile	Gly	Lys	Arg	Val	Asp	Lys	Thr	Phe	Leu	Pro	Ser	Leu	Ala	Ile
285			195					200					205			
286																
287	Ile	Ser	Leu	Glu	Asn	Ser	Trp	Ser	Ala	Leu	Ser	Lys	Gln	Ile	Gln	Ile
288		210					215					220				
289																
290		Ser	Thr	Asn	Asn	Gly	Gln	Phe	Glu	Thr		Val	Val	Leu	Ile	Asn
291	225					230					235					240
292			_		_	•					_		_	_	_	_
293	Ala	GIn	Asn	Gln		Val	Met	Ile	Thr		Val	Asp	Ala	Gly		Val
294					245					250					255	
295 296	mb	a	<b>3</b>	<b>T</b> 1 ~	<b>.</b> 1.	T	<b>.</b>	<b>T</b>	•	•	•	•				
297	Thr	ser	ABN	260	ATA	Leu	гел	ren		Arg	Asn	Asn	Met		Ата	wet
298				200					265					270		
299	Aan	Aan	Agn	17 n 1	Dro	Met	Thr	Gln.	g.~	Dho	C1	Cara.	C1	C	The seas	210
300	Aop	reb	275	val	PIO	Mec	1111	280	261	FIIG	GIA	Сув	285	ser	TÄT	AId
301			2,5					200					203			
302	Ile															
303																
304																
305	(2)	IN	ORM	ATION	I FOI	R SE	O ID	NO:	5:							
306	- •						_									
307	į)	L)	SI	EQUE	NCE (	CHAR	ACTE	RIST	cs:							
308			(2	4) I	LENG:	гн: 🤇	452	base	pa:	rs	> r	- Q ·	~ ~	hed	K	numbering
309			(1	3) 7	CYPE:	: nuc	cleid	ac:	ld-		•		_		_ , `	178
310			((	2) 8	STRAI	NDEDI	NESS:	: si	ngle							,
311			(I	) 1	ropo1	LOGY	: li:	near								
312																

### Raw Sequence Listing

03/29/93 15:40:37 S543.raw

313	(ii)	MOLECULE TYPE: DNA	(genomic)	
314 315	(iii)	HYPOTHETICAL: NO		
316	(111)	miroliibiican. No		
317	(iv)	ANTI-SENSE: NO		
318				
319	(vi)	ORIGINAL SOURCE:		
320		(A) ORGANISM: Ory:	a sativa	
321 322	(vii)	IMMEDIATE SOURCE:		
323	( \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	(B) CLONE: alpha-a	ໜໄລຊວ	
324		(), 0201.21 0251.0	.,,	
325	(ix)	FEATURE:		
326			(B) LOCATION: 121316	
327		(B) LOCATION: 12.	. 1316	
328 329	(xi)	CPOHENCE DECORTORS	I. CEO ID NO. E.	
330	(YT)	SEQUENCE DESCRIPTION	A: SEA IN WA: 2:	
331	CCTCGAG	STG C ATG CAG GTG CTG	AAC ACC ATG GTG AAC A CAC TTC T	rg (50) ONLY
332			2 22 22 22 22 22 22 22 22 22 22 22 22 2	
333			Asn Thr Met Val Asn Lys His Phe	Leu Ha are
334		1	5 10	on this
335 336	TOO OMM	MOG CING CING AMG CING		
337	TCC CTT	TCG GTC CTC ATC GTC (	C CTT GGC CTC TCC TCC AAC TTG	ACA 98 \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
338	Ser Leu	Ser Val Leu Ile Val I	eu Leu Gly Leu Ser Ser Asn Leu 3	on this  on this  in the Re-calwing  numbering
339	15	20	25	ice carulate
340				numbering
341	GCC GGG	CAA GTC CTG TTT CAG	GA TTC AAC TGG GAG TCG TGG AAG (	GAG 146
342 343	31a Gl.	Glm Wal You Pho Glm (		
344	30	35	Ly Phe Asn Trp Glu Ser Trp Lys ( 40	31u 45
345	30	33	40	
346	AAT GGC	GGG TGG TAC AAC TTC	G ATG GGC AAG GTG GAC GAC ATC	SCC 194
347				
348	Asn Gly		eu Met Gly Lys Val Asp Asp Ile A	Ala
349 350		50	55 60	
351	GCA GCC	GGC ATC ACC CAC GTC 1	GG CTC CCT CCG CCG TCT CAC TCT (	FTC 242
352		dde are ace eae die i	ed cic cci ccd ccd ici cac ici c	242
353	Ala Ala	Gly Ile Thr His Val 7	p Leu Pro Pro Pro Ser His Ser V	/al
354		65	70 75	
355				
356 357	GGC GAG	CAA GGC TAC ATG CCT (	GG CGG CTG TAC GAT CTG GAC GCG T	TCT 290
35 <i>7</i> 358	Glv Glu	Gin Gly Tyr Met Pro	y Arg Leu Tyr Asp Leu Asp Ala S	lor
359	or, ora	80	so s	961
360				
361	AAG TAC	GGC AAC GAG GCG CAG	C AAG TCG CTG ATC GAG GCG TTC	CAT 338
362	T	al al al-	• •. • • • • • • • • • • • • •	
363 364	Lys Tyr 95	Gly Asn Glu Ala Gln I 100	eu Lys Ser Leu Ile Glu Ala Phe I	lis
20 <del>4</del>	73	100	105	

### Raw Sequence Listing

03/29/93 15:40:44 S543.raw

365																	
366	GGC	AAG	GGC	GTC	CAG	GTG	ATC	GCC	GAC	ATC	GTC	ATC	AAC	CAC	CGC	ACG	386
367																	
368	Gly	Lys	Gly	Val	Gln	Val	Ile	Ala	Asp	Ile	Val	Ile	Asn	His	Arq	Thr	
369	110		_			115			_		120				•	125	
370																	
371	GCG	GAG	CAC	AAG	GAC	GGC	CGC	GGC	ATC	TAC	TGC	СТС	ттс	GAG	GGC	GGG	434
372								-					***	ong	GGC	333	474
373	Δla	Glu	Hie	Tara	Agn	Gly	Ara	Gly	710	Th. 2	Cara	Lou	Pho	G1	<b>C1</b>	C1	
374		014	*****	Ly 5	130	GIY	ALG	Gry	116	135	Cyb	пец	FIIG	GIU	–	GIA	
375					130					133					140		
376	3.00	ccc	CAC	maa	000	ama	~~	maa	-	-	a. a						400
377	ACG	CCC	GAC	100	CGC	CTC	GAC	TGG	GGC	CCG	CAC	ATG	ATC	TGC	CGC	GAC	482
	m1	D	•	<b>a</b>					~3	_	1			_	_	_	
378	Inr	PIO	Авр		Arg	Leu	Asp	Trp		Pro	His	Met	Ile	_	Arg	Asp	
379				145					150					155			
380																	
381	GAC	CCC	TAC	GGC	CAT	GGC	ACC	GGC	AAC	CCG	GAC	ACC	GGC	GCC	GAC	TTC	530
382																	
383	Asp	Pro		Gly	Asp	Gly	Thr	Gly	Asn	Pro	Asp	Thr	Gly	Ala	Asp	Phe	
384			160					165					170				
385																	
386	GCC	GCC	GCG	CCG	GAC	ATC	GAC	CAC	CTC	AAC	AAG	CGC	GTC	CAG	CGG	GAG	578
387																•	
388	Ala	Ala	Ala	Pro	Asp	Ile	Asp	His	Leu	Asn	Lys	Arg	Val	Gln	Arg	Glu	
389		175					180					185					
390																	
391	CTC	ATT	GGC	TGG	CTC	GAC	TGG	CTC	AAG	ATG	GAC	ATC	GGC	TTC	GAC	GCG	626
392																	
393	Leu	Ile	Gly	Trp	Leu	Asp	Trp	Leu	Lys	Met	Asp	Ile	Gly	Phe	Asp	Ala	
394	190					195					200		_		_	205	
395																	
396	TGG	CGC	CTC	GAC	TTC	GCC	AAG	GGC	TAC	TCC	GCC	GAC	ATG	GCA	AAC	ATC	674
397																	
398	Trp	Arg	Leu	Asp	Phe	Ala	Lvs	Glv	Tvr	Ser	Ala	qaA	Met	Ala	Lvs	Ile	
399	-	_		_	210		-	•	•	215					220		
400																	
401	TAC	ATC	GAC	GCC	ACC	GAG	CCG	AGC	TTC	GCC	GTG	CCC	GAG	АТА	TCG	ACG	722
402																	
403	Tvr	Ile	asp	Ala	Thr	Glu	Pro	Ser	Phe	Ala	Val	λla	Glu	Tle	Trn	Thr	
404	-3-			225					230					235			
405														233			
406	TCC	ATG	GCG	AAC	GGC	GGG	GAC	GGC	AAG	cca	AAC	TAC	GAC	CAG	AAC	GCG	770
407					-	-	Onc	000	ALC	CCG	MC	IAC	GAC	CAG	AAC	GCG	,,,
408	Ser	Met	aΙΔ	λen	Glv	Gly	) an	Glv	Tara	Pro	Agn	Tarr	Agn	Gln.	Acn	<b>7</b> 15	
409	Del	Mec	240	AOII	GIY	Gry	veħ	245	пув	110	ADII	TAT	250	GIII	WRII	AIG	
410			240					243					250				
411	ראת	CGC	CAC	GNC	СТС	GTC	224	TOO	ama	(I) III	CCM	ama	aaa	000	000	33C	010
412	CAC	<b>-3</b> -3-3	CAG	GAG	CIG	GIC	MAC	196	GIC	GAI	CGI	GIC	GGC	GGC	GCC	MAC	818
413	ui.	λ <b>~~</b> ~	<b>G</b> 1∽	G1	Lev	37 o 3	<b>7</b> ~~	Ф	77 n 7	<b>3</b>	<b>3</b>	77c 7	01	<b>41</b>		3	
414	итя		GIII	GIU	пел	Val		rrb	val	vab	Arg		стА	стА	WIS	ABII	
		255					260					265					
415																	
416	300	AAC	000	3 ~~	~~~	m=-	a = ~	me~									866

### Raw Sequence Listing

03/29/93 15:40:50 S543.raw

417																	
418	Ser	Asn	Gly	Thr	Ala	Phe	Asp	Phe	Thr	Thr	Lva	Glv	Tla	T.011	λan	Val	
419	270		2			275					280	<b>-</b> -,				285	
420																	
421	GCC	GTG	GAG	GGC	GAG	CTG	TGG	CGC	CTC	CGC	GGC	GAG	GAC	GGC	AAG	GCG	914
422																	
423	Ala	Val	Glu	Gly	Glu	Leu	Trp	Arg	Leu	Arg	Gly	Glu	Asp	Gly	Lys	Ala	
424					290					295					300		
425																	
426 427	ccc	GGC	ATG	ATC	GGG	TGC	TGG	CCG	GCC	AAG	GCG	ACG	ACC	TTC	GTC	GAC	962
428	Dro	<b>~1</b>	Wat	T1.	~1··	W	///	D===	31.	T	31.	m1	m1	nh.	**- 1	3	
429	PIO	GIY	Met	305	GIY	пр	пр	Pro	310	гЛя	Ата	Thr	Thr	315	vaı	Asp	
430				505					310					313			
431	AAC	CAC	GAC	ACC	GGC	TCG	ACG	CAG	CAC	CTG	TGG	CCG	TTC	CCC	TCC	GAC	1010
432																	
433	Asn	His	Asp	Thr	Gly	Ser	Thr	Gln	His	Leu	Trp	Pro	Phe	Pro	Ser	Asp	
434			320					325					330				
435																	
436	AAG	GTC	ATG	CAG	GGC	TAC	GCA	TAC	ATC	CTC	ACC	CAC	CCC	GGC	AAC	CCA	1058
437 438	Tara	1707	Woh		<b>~1</b>	TT	31-	m	<b>T</b> 1.	T	m\	***	D	<b>~1</b>	<b>3</b>	D	
439	гуя	335	Met	GIN	GIY	TYF	340	туг	TTE	гел	Thr	345	Pro	GIY	Asn	Pro	
440		333					340					343					
441	TGC	ATC	TTG	TAC	GAC	CAT	TTC	TTC	GAT	TGG	GGT	CTC	AAG	GAG	GAG	ATC	1106
442																	
443	Сув	Ile	Phe	Tyr	Asp	His	Phe	Phe	Asp	Trp	Gly	Leu	Lys	Glu	Glu	Ile	
444	350					355					360					365	
445																	
446	GAG	CGC	CTG	GTG	TCA	ATC	AGA	AAC	CGG	CAG	GGG	ATC	CAC	CCG	GCG	AGC	1154
447 448	<b>~1</b>	3	T	77.7	g	<b>71</b> -	<b>3</b>	<b>.</b>	<b>3</b>	<b>41</b>	<b>a</b> 1	<b>-1</b> -	••• ·	<b>D</b>		<b>a</b>	
449	GIU	Arg	Leu	vai	370	TIE	Arg	ABN	Arg	375	GIA	TTE	HIS	PFO	380	ser	
450					3,0					3,3					300		
451	GAG	CTG	CGC	ATC	ATG	GAA	GCT	GAC	AGC	GAT	CTC	TAC	CTC	GCG	GAG	ATC	1202
452																	
453	Glu	Leu	Arg	Ile	Met	Glu	Ala	qaA	Ser	Asp	Leu	Tyr	Leu	Ala	Glu	Ile	
454				385					390					395			
455																	
456	GAT	GGC	AAG	GTG	ATC	ACA	AAG	ATT	GGA	CCA	AGA	TAC	GAC	GTC	GAA	CAC	1250
457 458	7 ~~	C1	T	37a 1	T1.	Th-	7	T1.	<b>~1</b>	D===	3	TT	3	77-7	<b>a</b> 1	774 -	
459	мар	GIY	Lys 400	vai	116	Inr	гуя	405	GIY	PIO	Arg	TAL	410	val	GIU	HIS	
460			100										TIO				
461	CTC	ATC	CCC	GAA	GGC	TTC	CAG	GTC	GTC	GCG	CAC	GGT	GAT	GGC	TAC	GCA	1298
462																	
463	Leu		Pro	Glu	Gly	Phe		Val	Val	Ala	His	Gly	Asp	Gly	Tyr	Ala	
464		415					420					425					
465	3 m~	m~~	~- ~		<b>.</b>	ma			. ma-			~m~-	···		. ~		
466 467	ATC	TGG	GAG	AAA	ATC	TGA	∍CGC1	ACG 1	ATGA(	.'GAGI	AC T	CTCA	3TTT2	A GC	AGAT'	AATT	1353
467	Tle	Trn	Glu	Tare	T.T 🗢												
100	***	בבט	GIU	r J B	71.0												

### Raw Sequence Listing

03/29/93 15:40:57 S543.raw

469 470	430		43	5											
471 472 473	CCTGCGI	ATT TTA	CCCTGAC (	GGTATACG	T ATATAC	GTGC CGGC	AACGAG CTG	FATCCGA 1413							
474 475 476	TCCGAAT	TAC GGA	IGCAATT (	TCCACGAA	G TCCTCG	AGG		1452							
477 478 479	(2) IN	FORMATI	ON FOR SE	Q ID NO:	6:										
480 481	(i)		ENCE CHAP												
482		(A) (B)		434 amin											
483 484		(D) Topology: linear  (ii) MOLECULE TYPE: protein													
485	(ii)	(ii) MOLECULE TYPE: protein													
486 487	(xi) MOLECULE TYPE: protein  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:														
488	(XI)	SEQU.	ENCE DESC	RIPTION:	REG ID	NO: 6:									
489 490 491	Met Glr 1	val Le	u Asn Thr 5	Met Val	Asn Lys 10		Leu Ser Leu 19								
492 493 494	Val Leu	ı Ile Va 2		Gly Leu	Ser Ser 25	Asn Leu	Thr Ala Gly 30	y Gln							
495 496 497	Val Lev	Phe Gl:	n Gly Phe	Asn Trp 40		Trp Lys	Glu Asn Gly 45	y Gly							
498 499 500	Trp Tyr		e Leu Met	Gly Lys 55	Val Asp	Asp Ile 60	Ala Ala Ala	a Gly							
501 502 503	Ile Thr	His Va	l Trp Leu 70		Pro Ser	His Ser	Val Gly Glu	ı Gln 80							
504 505 506	Gly Tyr	Met Pr	o Gly Arg 85	Leu Tyr	Asp Leu 90	_	Ser Lys Ty: 95	_							
507 508 509	Asn Glu	Ala Gli 10		Ser Leu	Ile Glu 105	Ala Phe	His Gly Lys 110	g Gly							
510 511 512	Val Glr	Val Il	e Ala Asp	Ile Val 120		-	Thr Ala Glu 125	ı His							
513 514 515	Lys Asp 130		g Gly Ile	Tyr Cys 135	Leu Phe	Glu Gly 140	Gly Thr Pro	Asp							
516 517 518	Ser Arg	Leu As	p Trp Gly 150		Met Ile	Cys Arg . 155	Asp Asp Pro	7 Tyr 160							
519 520	Gly Asp	Gly Th	r Gly Asr 165	Pro Asp	Thr Gly 170		Phe Ala Ala 175								

### Raw Sequence Listing

03/29/93 15:41:04 S543.raw

521																
522	Pro	qaA	Ile	Asp	His	Leu	Asn	Lvs	Arg	Val	Gln	Ara	Glu	T. <b>-</b> 11	Tla	Gly
523				180				-,-	185			9		190		GLY
524																
525	Trp	Leu	Asp	qrT	Leu	Lys	Met	Asp	Ile	Glv	Phe	Asp	Ala	Trp	Ara	Leu
526	_		195	-		•		200		•			205		3	
527																
528	Asp	Phe	Ala	Lys	Gly	Tyr	Ser	Ala	Asp	Met	Ala	Lys	Ile	Tyr	Ile	Asp
529	_	210		-	_	•	215		_			220		•		•
530																
531	Ala	Thr	Glu	Pro	Ser	Phe	Ala	Val	Ala	Glu	Ile	Trp	Thr	Ser	Met	Ala
532	225					230					235	_				240
533																
534	Asn	Gly	Gly	Asp	Gly	Lys	Pro	Asn	Tyr	Asp	Gln	Asn	Ala	His	Arg	Gln
535					245					250					255	
536																
537	Glu	Leu	Val	Asn	Trp	Val	Asp	Arg	Val	Gly	Gly	Ala	Asn	Ser	Asn	Gly
538				260					265					270		_
539																
540	Thr	Ala	Phe	Asp	Phe	Thr	Thr	Lys	Gly	Ile	Leu	Asn	Val	Ala	Val	Glu
541			275					280					285			
542																
543	Gly		Leu	Trp	Arg	Leu	Arg	Gly	Glu	Asp	Gly	Lys	Ala	Pro	Gly	Met
544		290					295					300				
545																
546		Gly	Trp	Trp	Pro		Lys	Ala	Thr	Thr		Val	Asp	Asn	His	Asp
547	305					310					315					320
548										_						
549	Thr	Gly	Ser	Thr		His	Leu	Trp	Pro		Pro	Ser	Asp	Lys		Met
550					325					330					335	
551	<b>41</b>	~7 -	_		_		_	_,	•	_		_	_	_		
552 553	GIN	GIA	Tyr		ıyr	тте	Leu	Thr		Pro	GТĀ	Asn	Pro	_	Ile	Phe
554				340					345					350		
555	TT	3	77.3	Db.	nh.	<b>3</b>	m	<b>a</b> 1		<b>.</b>	<b>~</b> 1	<b>~</b> 1	-1-	<b>a</b> 1	•	
556	ıyı	Авр	His 355	Pne	Pne	Авр	Trp	_	ren	гля	GIU	GIU		GIU	Arg	Leu
557			333					360					365			
558	77-3	C.~	Ile	λ <b></b>	X am	λ	<b>63</b> m	C3	T3.0	TT -	D===	31.	g	<b>a</b> 1	T	3
559	Val	370	116	Arg	ASII	Arg	375	GIY	116	urs	PIO	380	ser	GIU	ьеu	Arg
560		370					3/3					300				
561	Tla	Mat	Glu	λla	A an	Cor	Nan	Lou	There	Lou	λl a	C1	т1.	N an	G1	*
562	385	Mec	GIU	ALG	veb	390	veħ	пец	TYT	пеа	395	GIU	116	Asp	GIY	400
563	505					330					333					400
564	Val	Tle	Thr	Lvs	Tle	Glv	Pro	Ara	Tur	λan	Val	Gl 11	Hie	T.011	Tla	Pro
565				-10	405	~-1		9	-1-	410	<b>741</b>	<b>314</b>	***	Leu	415	-10
566																
567	Glu	Glv	Phe	Gln	Val	Val	Ala	His	Glv	Asp	Glv	Tvr	Ala	IJe	Trp	Glu
568		3		420					425	<b>-</b>	1	-1-		430		
569																
570	Lys	Ile														
571	-															
572																

### Raw Sequence Listing

03/29/93 15:41:05 S543.raw

573	(2) INF	ORMATION FOR SEQ ID NO:7:	
574 575	(i)		
576	(1)	SEQUENCE CHARACTERISTICS:	
577		(A) LENGTH: '709 base pairs	
		(B) TYPE: nucleic acid	
578 579		(G) STRANDEDNESS: single	
580		(D) TOPOLOGY: linear	
581	(44)	MOLECULE TYPE: cDNA to mRNA	
582	(11)	MODECODE TIPE. COMA CO MININA	
583	(111)	HYPOTHETICAL: NO	
584	\/		
585	(iv)	ANTI-SENSE: NO	
586			
587	(vi)	ORIGINAL SOURCE:	
588		(A) ORGANISM: Homo sapiens	
589		<u>-</u>	
590	(vii)	IMMEDIATE SOURCE:	
591		(B) CLONE: alpha-hemoglobin	
592			
593	(ix)	FEATURE:	
594		(A) NAME/KEY: transit_peptide (B) LOCATION: 26241	
595		(B) LOCATION: 26241	
596			
597	(ix)	FEATURE:	
598		(A) NAME/KEY: CDS	
599		(B) LOCATION: 245670	
600 601	(no. i )	GEOGRAFIA DEGREE DE CONTRACTO DE LA CONTRACTO DE CONTRACT	
602	(XI)	SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
603	CTCGAGGG	CA TCTGATCTTT CAAGAATGGC ACAAATTAAC AACATGGCAC AAGGGATACA	
604	CICGAGGG	CR ICIGRICIII CAAGARIGGC ACAAAIIAAC AACAIGGCAC AAGGGATACA	60
605	AACCCTTA	AT CCCAATTCCA ATTTCCATAA ACCCCAAGTT CCTAAATCTT CAAGTTTTCT	120
606		Committee Miller Accountil Committee Cardillici	120
607	TGTTTTTG	GA TGTAAAAAAC TGAAAATTC AGCAAATTCT ATGTTGGTTT TGAAAAAAGA	180
608			
609	TTCAATTT	TT ATGCAAAAGT TTTGTTCCTT TAGGATTTCA GCAGGTGGTA GAGTTTCTTG	240
610			
611	CATG GTG	CTG TCT CCT GCC GAC AAG ACC AAC GTC AAG GCC GCC TGG GGC	289
612			
613	Val	Leu Ser Pro Ala Asp Lys Thr Asn Val Lys Ala Ala Trp Cly	
614	1	5 10 15	
615			
616	AAG GTT (	GGC GCG CAC GCT GGC GAG TAT GGT GCG GAG GCC CTG GAG AGG	337
617			
618	Lys Val (	Gly Ala His Ala Gly Glu Tyr Gly Ala Glu Ala Leu Glu Arg	
619		20 25 30	
620	3 mg mmg 4	CMC MCC MMC CCC 200 200 200 220 200 200 820 820 820 820	~ ~ -
621 622	AIG TTC (	CTG TCC TTC CCC ACC ACC AAG ACC TAC TTC CCG CAC TTC GAC	385
623	Met Phe 1	Leu Ser Phe Pro Thr Thr Lys Thr Tyr Phe Pro His Phe Asp	
624	Met File I	35 40 45	

### Raw Sequence Listing

03/29/93 15:41:17 S543.raw

625																	
626	CTG	AGC	CAC	GGC	тст	GCC	CAG	GTT	AAG	GGC	CAC	GGC	AAG	AAG	стс	GCC	433
627												-				000	
628	Leu	Ser	His	Gly	Ser	Ala	Gln	Val	Lys	Gly	His	Gly	Lys	Lys	Val	Ala	
629			50	-				55	•	•		•	60	•			
630																	
631	GAC	GCG	CTG	ACC	AAC	GCC	GTG	GCG	CAC	GTG	GAC	GAC	ATG	CCC	AAC	GCG	481
632																	
633	Asp	Ala	Leu	Thr	Asn	Ala	Val	Ala	His	Val	Asp	Asp	Met	Pro	Asn	Ala	
634		65					70					75					
635																	
636	CTG	TCC	GCC	CTG	AGC	GAC	CTG	CAC	GCG	CAC	AAG	CTT	CGG	GTG	GAC	CCG	529
637	_	_		_	_	_	_			•				_			
638		Ser	Ala	Leu	Ser		Leu	His	Ala	His		Leu	Arg	Val	Asp		
639 640	80					85					90					95	
641	CTC	220	TO CO	220	CTC.	CITIA	3.00	as a	maa	ama	ama	ama	3.00	ama	~~~	aaa	500
642	GIC	AAC	110	AAG	CIC	CIA	AGC	CAC	TGC	CIG	CTG	GIG	ACC	CTG	GCC	GCC	577
643	٧al	λan	Pho	Lys	T.611	T.011	Ser	Hio	Cva	T.O.	T. 011	77 n 7	Thr	Lou	21-	77-	
644		*****		ny 5	100	100	Der	1110	Cyb	105	пец	Val	1111	пеа	110	AIG	
645																	
646	CAC	CTC	CCC	GCC	GAG	TTC	ACC	CCT	GCG	GTG	CAC	GCC	TCC	CTG	GAC	AAG	625
647																	
648	His	Leu	Pro	Ala	Glu	Phe	Thr	Pro	Ala	Val	His	Ala	Ser	Leu	Asp	Lys	
649				115					120					125		_	
650																	
651	TTC	CTG	GCT	TCT	GTG	AGC	ACC	GTG	CTG	ACC	TCC	AAA	TAC	CGT	TAA	GCTGGAG	677
652																	
653	Phe	Leu		Ser	Val	Ser	Thr		Leu	Thr	Ser	Lys	Tyr	Arg			
654			130					135					140				
655													,				
656 657	COM	777m		CGTT		am a	7000	3maa:			1	,	$^{\circ}$ G	1	$   \Delta U $	C10 (	number
658	CCI	-GG11	AGC (	CGII	-616	CI G	٠٠٠٠	3TCG/	A CC		_	-			. –		
659													ì	5		requ	c numberined at
660	(2)	IN	FORM	ATIO	V FOI	R SEC	о тр	NO :	R •					L.)			- 17-17-2
661	\-,						2		•				1	- المرا		~ d	L 06 77.5
662	(:	i)	SI	EQUE	NCE (	CHAR	ACTE	RIST:	ICS:					ſ	no		
663	•	•		.~	LENG:					acida	3			ŀ	\ \ \ \ \	- •	
664			(1	B) :	TYPE:	: aı	nino	acio	£								
665			(1	D) :	ropoi	LOGY	: 1:	inear	r								
666							,										
667	(i:	i)	M	OLEC	JLE :	TYPE	: pro	otei	a								
668																	
669																	
670	(i:	K)	S	EQUE	NCE I	DESCI	RIPT:	ON:	SEQ	ID 1	8:07	:					
671	49 - T	<b>.</b>		<b>n</b>		•	<b>-</b>	m1-			_			-	<b>~</b> -	-	
672		Leu	ser	Pro		Asp	Lys	Thr	Asn		Lys	Ala	Ala	Trp		Lys	
673 674	1				5					10					15		
675	Val	Glv	Ala	His	Alα	Glv	Glu	የ	Gl v	Als	GI 11	Δls	Len	G1 11	A~~	Met	
676		1		20				-1-	25	n. L. Cl	QI U	WT CL	Leu	30	nr 9	-46 ¢	
- · •														50			

### Raw Sequence Listing

03/29/93 15:41:24 S543.raw

677																	
678	Phe	Leu	Ser	Phe	Pro	Thr	Thr	Lys	Thr	Tyr	Phe	Pro	His	Phe	Asp	Leu	
679			35					40		-			45		_		
680																	
681	Ser	His	Gly	Ser	Ala	Gln	Val	Lys	Gly	His	Gly	Lys	Lys	Val	Ala	Asp	
682		50					55	_	_		_	60	_			-	
683																	
684	Ala	Leu	Thr	Asn	Ala	Val	Ala	His	Val	Asp	Asp	Met	Pro	Asn	Ala	Leu	
685	65					70				•	75					80	
686																	
687	Ser	Ala	Leu	Ser	Asp	Leu	His	Ala	His	Lys	Leu	Arg	Val	Asp	Pro	Val	
688					85					90		•		-	95		
689																	
690	Asn	Phe	Lys	Leu	Leu	Ser	His	Сув	Leu	Leu	Val	Thr	Leu	Ala	Ala	His	
691			_	100				•	105					110			
692																	
693	Leu	Pro	Ala	Glu	Phe	Thr	Pro	Ala	Val	His	Ala	Ser	Leu	qsA	Lys	Phe	
694			115					120					125	-	•		
695																	
696	Leu	Ala	Ser	Val	Ser	Thr	Val	Leu	Thr	Ser	Lys	Tyr	Arg				
697		130					135				_	140	Ĭ				
698																	
699																	
700	(2)	INF	ORMA'	rion	FOR	SEQ	ID 1	10:9	:								
701																	
702	i)	L)		EQUEI						•	_						,
703				A) L		_	-	-	-	8	, ?	Se	ج_	77-6	とべこ	+ page	
704			(1	B) T	YPE:	nuc.	leic	aci	2								
705				c) s:					gle								
706			(1	D) T(	OPOL	OGY:	line	ear									
707																	
708	(ii	L)	M	OLEC	JLE :	CYPE	: cD1	NA to	o mRI	AV							
709																	
710	(iii	L)	H	YPOTI	HETI(	CAL:	ИО										
711							_										
712	(iz	7)	Al	NTI-	SENSI	3: N	)										
713																	
714	iv)	L)		RIGII													
715			(2	A.) (	ORGAI	NISM	: Hoi	no sa	apie	ns							
716	,																
717	(vii	L)		MMED:													
718			(1	B) (	CLONI	s: D	eta-l	nemo	TOD:	ın							
719	/ 4 -	_\	979.1	- <b>3</b> mrs													
720	(i)	c)		EATU		/							<b></b>				
721										ptla	9 (B	) TO	CATI	ON:	26.	. 241	
722			()	B) 1	LOCA	LION	: 26	24.	L								
723 724	/2-	٠١	<b>273</b> 1	EATU	. ac												
72 <del>4</del> 725	(i)	<b>~</b> /			KE: NAME	/ppv		2									
725 726			-	•	LOCA:			-	0 5								
727			()	ا رو	LUCA.	TON	. 24:	٠٠٠٠)	55								
728	(xi		Q1	EQUE	י שיטע	ישפארי	יים ד	ron.	SEV.	י חד	ντ <b>Ω</b> •	٥.					
	/	-,	9,	_ ** ~ ~ ~ ,					~~~			- •					

### Raw Sequence Listing

03/29/93 15:41:31 S543.raw

729			
730	CTCGAGGGA TCTGATCTTT CAAGAATGGC	ACAAATTAAC AACATGGCAC AAGGGATACA	60
731	·		
732	AACCCTTAAT CCCAATTCCA ATTTCCATAA	ACCCCAAGTT CCTAAATCTT CAAGTTTTCT	120
733			220
734	TGTTTTTGGA TCTAAAAAAC TGAAAATTTC	AGCAAATTCT ATGTTGGTTT TGAAAAAAGA	180
735	101111100M 1CIMBBERG IGMEDINIIC	AGCARATICI RIGIIGGIII IGAAAAAAGA	180
736		MAGGAMMAGA GGAGGAGGAA GAGMMAGAMAG	242
737	IICARIIIII AIGCAAAAGI IIIGIICCII	TAGGATTTCA GCAGGTGGTA GAGTTTCTTG	240
738	CAMC COO CAC COO ACT COO CAC	330 mg	
	GATG GTG CAC CTG ACT CCT GAG GAG	AAG TOT GOO GIT ACT GOO CIG TGG	289
739			
740	Val His Leu Thr Pro Glu Glu	Lys Ser Ala Val Thr Ala Leu Trp	
741	1 5	10 15	
742			
743	GGC AAG GTG AAC GTG GAT GAA GTT	GGT GGT GAG GCC CTG GGC AGG CTG	337
744			
745	Gly Lys Val Asn Val Asp Glu Val	Gly Gly Glu Ala Leu Gly Arg Leu	
746	20	25 30	
747			
748	CTG GTG GTC TAC CCT TGG ACC CAG	AGG TTC TTT GAG TCC TTT GGG GAT	385
749			
750	Leu Val Val Tyr Pro Trp Thr Gln	Arg Phe Phe Glu Ser Phe Gly Asn	
751	35	40 45	
752		10	
753	CTG TCC ACT CCT GAT GCT GTT ATG	CCC AAC CCT AAC CTC AAC CCT CAT	433
754	cro ice aci cci gai gci gii aig	GGC AAC CCI AAG GIG AAG GCI CAI	433
755	Leu Ser Thr Pro Asp Ala Val Met	Clu Nam Dwe Ive Wel Ive Nie Wie	
756			
757	50 55	60	
		100 010 000 000 000 000 000	
758	GGC AAG AAA GTG CTG GGT GCC TTT	AGT GAT GGC CTG GCT CAC CTG GAC	481
759	al		
760	Gly Lys Lys Val Leu Gly Ala Phe		
761	65 70 IL	Canestra?	
762			(529) S 3 O
763	AAC CTC AAG GGC ACC TTT GCC ACCA	CTG AGT GAG CTG CAC TGT GAC AAG	(529)
764	•		
765	Asn Leu Lys Gly Thr Phe Ala Thr	Leu Ser Glu Leu His Cys Asp Lys	
766	80 85	90 95	
767			
768	CTG CAC GTG GAT CCT GAG AGC TTC	AGG CTC CTA GGC AAC GTG CTG GTC	577
769			
770	Leu His Val Asp Pro Glu Ser Phe	Arg Leu Leu Gly Asn Val Leu Val	
771	100	105 110	
772			
773	TGT GTG CTG GCG CAT CAC TTT GGC	AAA GAA TTC ACC CCA CCA GTG CAG	625
774			
775	Cys Val Leu Ala His His Phe Gly	Lys Glu Phe Thr Pro Pro Val Gln	
776		120 125	
777	<del></del>		
778	GCT GCC TAT CAG AAA GTG GTG GCT	GGT GTG GCT AAT GCC CTG GCC CAC	673
779			~ · <del>~</del>
780	Ala Ala Tyr Gln Lys Val Val Ala	Gly Val Ala Agn Ala Lon Ala Big	
. 50	jr ozm njo vat vat Ata	orl age were you are ned ate ute	

### Raw Sequence Listing

03/29/93 15:41:38 S543.raw

781 782		130		135	140	
783	AAG TAT	CAC TAA	GCTCGCT TI	CTTGCTGT C	CAATTTCTA TTAAAGG	TTC 722
784 785	T	*** _				
786	Lys Tyr 145					
787	110					
788	CTTTGTGGGG TCGAGGTCGA C					
789						
790						
791	(0)					
792 793	(2) IN	FORMATIO	N FOR SEQ	ID NO: 10:		
794	(i)	SROTTE	NCE CHARAC	CTERISTICS:		
795	(-/	_		amino aci	ds	
796			TYPE: amin			
797		(D)	TOPOLOGY:	linear		
798	(ii)	MOLEC	ULE TYPE:	protein		
799						
800	(xi)	SEQUE	NCE DESCRI	PTION: SEQ	ID NO: 10:	
801 802	Val Wie	Leu Thr	Pro Glu G	ilu Twa Som	Ala Val Thr Ala	I ou Tron Clas
803	1	neu IIII	5	erd mys ser	10	15
804	_		•			13
805	Lys Val	Asn Val	Asp Glu V	al Gly Gly	Glu Ala Leu Gly	Arg Leu Leu
806		20		25		30
807			_	_		
808	Val Val		Trp Thr G		Phe Glu Ser Phe	Gly Asp Leu
809 810		35		40	45	
811	Ser Thr	Pro Asp	Ala Val M	Set Glv Asn	Pro Lys Val Lys	Ala His Glv
812	50			55	60	
813						
814		Val Leu		he Ser Asp	Gly Leu Ala His	<del>-</del>
815	65		70		75	80
816 817	Lou Iva	Clar The	Dho Nio T	The Tou Com	Glu Leu His Cys	Nam Your You
818	ned nys	GIY IIII	85	.nr heu ser	90	явр шув шец 95
819						33
820	His Val	Asp Pro	Glu Ser P	he Arg Leu	Leu Gly Asn Val	Leu Val Cys
821		100		105	_	110
822						
823	Val Leu		His Phe G		Phe Thr Pro Pro	Val Gln Ala
824 825		115		120	125	
826	Ala Tvr	Gln Lvs	Val Val A	la Glv Val	Ala Asn Ala Leu	Ala His Lve
827	130			35	140	
828			_			
829	Tyr His					
830	145					
831 832						
034						

### Raw Sequence Listing

03/29/93 15:41:44 S543.raw

833	(2)	INFORMATION FOR SEQ ID NO:11:
834		
835	(i)	SEQUENCE CHARACTERISTICS:
836		(A) LENGTH: 17 amino acids
837		(B) TYPE: amino acid
838		(D) TOPOLOGY: linear
839		
840	(ii)	MOLECULE TYPE: peptide
841		
842	(v)	FRAGMENT TYPE: N-terminal
843		1
844	(vi)	ORIGINAL SOURCE:
845		(A) ORGANISM: alkalophilic Bacillus sp.
846		(B) STRAIN: 38-2
847		
848	(vii)	IMMEDIATE SOURCE:
849		(B) CLONE: beta-cyclodextrin
850		
851	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 11) +his is an I 1
852		$\mathcal{O}$
853	Ala P	ro Asp Thr Ser Val Ser Asp Lys Glm Asp Phe Ser Thr Asp Val
854	1	5 10 15 letter
855		
856	Ile	

PAGE: 1

SEQUENCE VERIFICATION REPORT PATENT APPLICATION US/07/923,692

DATE: 03/29/93 TIME: 15:41:46

S543

LINE ERROR

ORIGINAL TEXT

248	Wrong Amino Acid Designator	Met lle Arg Phe Leu Val Leu Ser Leu Leu
245	Entered and Calc. Seq. Length differ	(x1) / SEQUENCE DESCRIPTION: SEQ ID NO
329	Entered and Calc. Seq. Length differ	(xi) SEQUENCE DESCRIPTION: SEQ ID NO
601		(xi) SEQUENCE DESCRIPTION: SEQ ID NO
728	Entered and Calc. Seq. Length differ	(xi) SEQUENCE DESCRIPTION: SEQ ID NO
851	Wrong Sequence Number	(xi) SEQUENCE DESCRIPTION: SEQ ID NO

PAGE: 1

SEQUENCE MISSING ITEM REPORT PATENT APPLICATION US/07/923,692

DATE: 03/29/93 TIME: 15:41:46

S543

MANDATORY IDENTIFIER THAT WAS NOT FOUND

COUNTRY

PRIOR APPLICATION DATA More Identifiers Found Than Max Allowed

PAGE:

1

SEQUENCE CORRECTION REPORT PATENT APPLICATION US/07/923,692

DATE: 03/29/93 TIME: 15:41:46

S543

LINE ORIGINAL TEXT

CORRECTED TEXT

851 (xi) SEQUENCE DESCRIPTION: SEQ ID NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: IL

BACKUP/RESTORE TRANSACTION

Transaction Number:

933

03/29/93

START TIME: 15:22:19 END TIME: 15:42:01

PROCESSING TIME: 00:19:42 Input Set: S543

DOSCOPY

US/07/923,692

Translog Code Application Serial Number National PCT N

Admendment

Application Class Application File Dt. 435

19920731